Journal of Computational and Applied Mathematics 38 (1991) 323–334
North-Holland

Solving nonlinear parabolic problems with result verification. Part I: One-space dimensional case

Mitsuhiro T. Nakao

Department of Mathematics, Faculty of Science, Kyushu University 33, Fukuoka 812, Japan

Received 30 July 1990
Revised 13 May 1991

Abstract

Nakao, M.T., Solving nonlinear parabolic problems with result verification. Part I: One-space dimensional case, Journal of Computational and Applied Mathematics 38 (1991) 323–334.

We propose some numerical methods for the automatic proof of existence of weak solutions for parabolic initial boundary value problems with one space dimension. It also means that one can obtain a posteriori error bounds for the approximate solutions of the problems. Based upon Schauder's fixed-point theorem, a verification condition is formulated and, by the use of finite-element approximation and its error estimates for a simple parabolic problem, we present a numerical verification algorithm of exact solutions in a computer. Some numerical examples which are verified by the method are illustrated.

Keywords: Parabolic problem, finite-element method, error estimates, fixed-point theorem.

1. Introduction

In recent years, various numerical methods with guaranteed error bounds, utilizing the interval analysis as the main tool, have been developed [6–8,18]. Further, similar attempts are also done for functional equations such as ordinary differential equations, integral equations and special functional equations [3–5,9,17,19]. These methods are distinguished by the fact that the existence of the exact solution for the original problem can be verified in the process of calculation of approximation itself, even if they are a priori unknown, as well as guaranteed accuracy of the approximate solution. For partial differential equations, however, there are very few such approaches up to now except the author's own reports [13–16] in which some numerical verification methods for elliptic problems are considered. In this paper, we attempt an extension of these methods to nonlinear parabolic initial boundary value problems of one space dimension. Our verification technique may, in principle, also be applied to the multi-space dimensional case. We will treat such problems in a forthcoming paper.

In the following section, we formulate the solution of the parabolic problem as a fixed point for a compact operator, and the fundamental principle for verification is presented. And in Section 3, using the finite-element approximation and the error estimates for a basic linear parabolic problem, the verification conditions in computer are derived. These conditions are clarified for both continuous time, i.e., semidiscrete finite-element approximation, and full discretization in space and time. Only for the case of full discretization, in the final section, we show a verification algorithm using the iterative procedure in computer and illustrate some numerical examples.

2. Problem and fixed-point formulation

Consider the following one-space dimensional nonlinear parabolic problem:

$$\begin{cases} u_t - u_{xx} = f(x, t, u), & (x, t) \in \Omega \times J, \\ u(x, 0) = 0, & x \in \Omega, \\ u(x, t) = 0, & (x, t) \in \partial\Omega \times J, \end{cases} \tag{2.1}$$

where Ω is a bounded open interval on $\mathbb{R}$. Let $J = (0, T)$ with $T > 0$, and let $Q = \Omega \times J$.

We denote by L^2 and H^m, H_0^m for any integer m, the usual Lebesgue and Sobolev spaces on Ω, respectively, and by $(\cdot, \cdot)_\Omega$, $\|\cdot\|_\Omega$ and $\|\cdot\|_m$ their natural inner products and norms, respectively. But, note that, for $m = 1$, we use $(\nabla u, \nabla v)_\Omega$ as the inner product and corresponding norm $\|u\|_1^2 = (\nabla u, \nabla u)_\Omega$ on H_0^1. Here, and also from now on, we use the notation $\nabla u \equiv u_x$. Also, we simply denote by $(\cdot, \cdot)$ and $\|\cdot\|$ the L^2 inner product and norm on Q, respectively. For nonnegative integers r and s, we define $H^r(J; H^s)$ to be the completion of $C^\infty(J; H^s)$ in the norm

$$\left(\|u\|_{H^r(J;H^s)}\right)^2 = \sum_{j=0}^{r} \int_J \left\| \left(\frac{\partial}{\partial t}\right)^j u(\cdot, t) \right\|_s^2 \mathrm{d}t,$$

where $C^\infty(J; H^s)$ denotes the set of infinitely differentiable functions from J into H^s for which all derivatives have continuous extensions to $\bar{J}$. Also let $H_0 \equiv L^2(J; H_0^1)$ and let $H \equiv H_0 \cap H^1(J; L^2)$. Then H is the Hilbert space with the following inner product [10]:

$$\langle u, v \rangle_Q \equiv \int_J (\nabla u, \nabla v)_\Omega \, \mathrm{d}t + \int_J (u_t, v_t)_\Omega \, \mathrm{d}t.$$

We now suppose the following conditions for the nonlinear map f in (2.1).

(A1) $f(\cdot, u) \in H^1(J; L^2)$ for any $u \in H \cap H^1(J; H^1)$.

(A2) For each bounded subset U of H, $f(\cdot, U) \equiv \{ f(\cdot, u) \mid u \in U \}$ is also bounded in $L^2(J; L^2)$.

(A3) f is the continuous map from any bounded set in H to $L^2(J; L^2)$ with the norm in H_0.

The typical example of f satisfying above conditions is

$$f(\cdot, y) = py^n + q,$$

or, in general, $f(\cdot, y) = \{\text{polynomial in } y\}$, where p and q are smooth functions in x and t. Note that such f is not continuous as the map from H_0 into $L^2(J; L^2)$, i.e., u^n does not belong to $L^2(J; L^2)$ for all $u \in H_0$ if $n > 1$.

Now as well known, for each $g \in L^2(J;\ L^2)$, the following parabolic problem has a unique solution ϕ in $H \cap L^2(J;\ H^2)$:

$$\begin{cases} \phi_t - \phi_{xx} = g, & (x,\ t) \in \Omega \times J, \\ \phi(x,\ 0) = 0, & x \in \Omega, \\ \phi(x,\ t) = 0, & (x,\ t) \in \partial\Omega \times J. \end{cases} \tag{2.2}$$

We denote the above correspondence by $\phi = Ag$. Further, we set $\tilde{H} = \{\phi \in H \mid \phi(\cdot,\ 0) = 0\}$, where $\phi(\cdot,\ 0) = 0$ means that $\lim_{t \to 0} \phi(\cdot,\ t) = 0$ in the $L^\infty(\Omega)$ sense. Then (2.2) can be written as the weak form: find $\phi \in \tilde{H}$ such that

$$(\phi_t,\ v)_\Omega + (\nabla\phi,\ \nabla v)_\Omega = (g,\ v)_\Omega, \quad v \in H_0^1,\ t \in J. \tag{2.3}$$

Therefore, we define a weak solution $u \in H$ for (2.1) as $u \in \tilde{H}$ satisfying

$$(u_t,\ v)_\Omega + (\nabla u,\ \nabla v)_\Omega = (f(\cdot,\ u),\ v)_\Omega, \quad v \in H_0^1,\ t \in J. \tag{2.4}$$

Then (2.4) can be rewritten in the fixed-point form:

$$u = Af(\cdot,\ u). \tag{2.5}$$

The following proposition provides us the fundamental principle of the verification.

Theorem 1. *For a bounded, convex, and nonempty subset $U \subset \tilde{H}$, if*

$$Af(\cdot,\ U) \subset U, \tag{2.6}$$

then there exists a solution $u \in \overline{U}$ for (2.5), where $\overline{U}$ means the closure of U in H.

Proof. First, we show that $\overline{U}$ is also closed in H_0. Now let $\overline{U}^0$ be the closure of $\overline{U}$ in H_0. When $u \in \overline{U}^0$ there exists a sequence such that $u_n \to u$ in H_0. Then, by the boundedness of $\overline{U}$ in H, there exists an element $\hat{u} \in H$ and a subsequence $\{u_{n_i}\}$ of $\{u_n\}$ such that

$$u_{n_i} \rightharpoonup \hat{u} \quad \text{(weakly) in } H.$$

Therefore, from the compactness of the injection $H \to L^2(J;\ L^2)$ [21],

$$u_{n_i} \to \hat{u} \quad \text{(strongly) in } L^2(J;\ L^2).$$

Further, naturally $u_n \to u$ in $L^2(J;\ L^2)$. Hence, we have $\hat{u} = u$. Thus from the weakly closedness of $\overline{U}$ (e.g., [22]), $u \in \overline{U}$ follows. Next, by the continuity of Af on H, we have

$$Af(\cdot,\ \overline{U}) \subset \overline{Af(\cdot,\ U)} \subset \overline{U}.$$

Further, using the well-known a priori estimates for the solution of (2.2) (e.g., [12]), it is seen that the previous defined map A is continuous from $L^2(J;\ L^2)$ into $H \cap L^2(J;\ H^2)$. Hence, by (A2) and the fact that the injection $H \cap L^2(J;\ H^2) \hookrightarrow H_0$ is compact, the composite map Af is also compact on $\overline{U}$ in the sense of H_0-norm. Therefore, taking notice of the convexity of $\overline{U}$, from Schauder's fixed-point theorem we obtain the desired conclusion. □

3. Rounding and verification conditions

For a given set $U \subset \tilde{H}$, since we cannot exactly calculate the left-hand side of (2.6), it is necessary, as in [12] etc., that, using some finite-dimensional spaces, we define an appropriate rounding operation from $\tilde{H}$ into them and estimate the rounding error.

3.1. Semidiscrete rounding

In this subsection, we consider a rounding only for the space, i.e., continuous in time. Let $\{S_h\}$ be a set of finite-dimensional subspaces of H_0^1 depending on h, $0 < h < 1$, with the following approximation property. For any $u \in H^2 \cap H_0^1$,

$$\inf_{\chi \in S_h} \| u - \chi \|_1 \leqslant C_1 h \, | u |_{H^2}, \tag{3.1}$$

where $| u |_{H^2}^2 \equiv \| u_{xx} \|_\Omega^2$ and C_1 can be numerically estimated as a positive real number.

Now let us define a subspace of $\tilde{H}$ as $H(J; S_h) \equiv \{ \phi \in \tilde{H} \mid \phi(t) \in S_h, \forall t \in J \}$. Then for each $g \in L^2(J; L^2)$ the semidiscrete rounding $R_x(Ag)$ of Ag into S_h is defined by $u_h \in H(J; S_h)$ satisfying

$$((u_h)_t, v)_\Omega + (\nabla u_h, \nabla v)_\Omega = (g, v)_\Omega, \quad v \in S_h, \ t \in J. \tag{3.2}$$

The following lemma provides the rounding error for $R_x(Ag)$.

Lemma 2. *For $g \in L^2(J; L^2)$, let $u = Ag$ and let u_h be the solution for* (3.2). *Then the following error estimates hold for the same constant C_1 as in* (3.1):

$$\left(\int_0^t \| u(s) - u_h(s) \|_1^2 \, \mathrm{d}s \right)^{1/2} \leqslant 2\sqrt{3} \, C_1 h \| g \|, \quad t \in J.$$

Proof. Setting $e = u - u_h$, from (2.3) and (3.2) we have

$$(e_t, v)_\Omega + (\nabla e, \nabla v)_\Omega = 0, \quad v \in S_h, \ t \in J.$$

Hence, for arbitrary $v \in S_h$ we have

$$\frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t} \| e \|_\Omega^2 + (\nabla e, \nabla e)_\Omega = (e_t, u - v)_\Omega + (\nabla e, \nabla (u - v))_\Omega. \tag{3.3}$$

Integrating both sides of (3.3) from 0 to t:

$$\int_0^t \| e \|_1^2 \, \mathrm{d}s \leqslant \| e_t \| \, \| u - v \| + \frac{1}{2} \left(\int_0^t \| e \|_1^2 \, \mathrm{d}s + \int_0^t \| u - v \|_1^2 \, \mathrm{d}s \right). \tag{3.4}$$

Using (3.1) and the well-known Aubin–Nitsche's trick (e.g., [1]), from (3.4) we obtain

$$\begin{aligned} \int_0^t \| e \|_1^2 \, \mathrm{d}s &\leqslant 2 \left(\| e_t \| C_1^2 h^2 \, | u |_{H^2(Q)} + \tfrac{1}{2} C_1^2 h^2 \, | u |_{H^2(Q)}^2 \right) \\ &\leqslant 2 C_1^2 h^2 \left\{ \left(\| u_t \| + \| (u_h)_t \| \right) | u |_{H^2(Q)} + \tfrac{1}{2} | u |_{H^2(Q)}^2 \right\}, \end{aligned} \tag{3.5}$$

where

$$| u |_{H^2(Q)}^2 \equiv \int_0^T | u(t) |_{H^2}^2 \, \mathrm{d}t.$$

Now, set $\phi = u$ and $v = u_t$ in (2.3) to obtain

$$\| u_t \|_\Omega^2 + \frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t} \| \nabla u \|_\Omega^2 \leqslant \tfrac{1}{2} \| u_t \|_\Omega^2 + \tfrac{1}{2} \| g \|_\Omega^2,$$

which implies

$$\| u_t \| \leqslant \| g \|. \tag{3.6}$$

Similarly, we have

$$\|(u_h)_t\| \leqslant \|g\|. \tag{3.7}$$

Further by $u_t = u_{xx} + g$,

$$|u|_{H^2(Q)} \leqslant \|u_{xx}\| \leqslant \|u_t\| + \|g\| \leqslant 2\|g\|. \tag{3.8}$$

Thus from (3.5)–(3.8) we obtain the desired estimates. □

Based upon the above lemma we now define the semidiscrete rounding error $\mathrm{RE}_x(Ag) \subset H$ as

$$\mathrm{RE}_x(Ag) = \left\{\phi \in \tilde{H} \mid \|\phi\|_{H_0} \leqslant 2\sqrt{3}\, C_1 h \|g\|\right\}.$$

Then, by Lemma 2 we have for any $g \in L^2(J;\ L^2)$,

$$Ag \in R_x(Ag) + \mathrm{RE}_x(Ag). \tag{3.9}$$

Furthermore, for the set of functions $G \subset L^2(J;\ L^2)$, we can similarly define the rounding $R_x(AG) \subset S_h$ and the rounding error $\mathrm{RE}_x(AG) \subset H$ as the union for each $g \in G$.

Now let $\mathbb{R}^+$ be the set of nonnegative real numbers. For any $\alpha \in \mathbb{R}^+$, let

$$[\alpha] \equiv \left\{\phi \in \tilde{H} \mid \|\phi\|_{H_0} \leqslant \alpha\right\}.$$

Also for a bounded, convex and closed subset U_h in $H(J;\ S_h)$ and $\alpha,\ \beta \in \mathbb{R}^+$, we define the ordered triple $(U_h,\ \alpha,\ \beta)$ as

$$(U_h,\ \alpha,\ \beta) = \left\{\phi \in \tilde{H} \mid \phi \in U_h + [\alpha] \text{ and } \|\phi_t\| \leqslant \beta\right\}.$$

Then by Theorem 1 and Lemma 2, we have the following verification condition.

Theorem 3. *Let $G = f(\cdot,\ U)$ for $U = (U_h,\ \alpha,\ \beta)$. If*

$$R_x(AG) \subset U_h, \tag{3.10}$$

$$\|\mathrm{RE}_x(AG)\|_{H_0} \leqslant \alpha, \tag{3.11}$$

$$\|G\| \leqslant \beta, \tag{3.12}$$

then there exists a solution $u \in \overline{U}$ for (2.5). Here, the norm for the set of functions implies the supremum for norms of all functions in the same set.

Remark 4. Note that we have used in the above theorem the fact that (3.12) implies $\|(Ag)_t\| \leqslant \beta$, $\forall g \in G$, by virtue of the estimate (3.6).

Of course, we cannot directly calculate the rounding $R_x(AG)$ for U. Let $\{\phi_j\}_{j=1,\ldots,M}$ be a basis of S_h and let $\tilde{U}_h = R_x(AG) \equiv \sum_{j=1}^{M} b_j(t)\phi_j(x)$; then $\boldsymbol{b} = (b_j(t))$ is an interval solution of the following system of ordinary differential equations with constant coefficients:

$$\frac{\mathrm{d}\boldsymbol{b}}{\mathrm{d}t} = P\boldsymbol{b} + F(t), \quad \boldsymbol{b}(0) = 0. \tag{3.13}$$

Thus we obtain the rounding $R_x(AG)$ provided that the solution for (3.13) can be obtained with guaranteed error bounds. Since, in general, (3.13) becomes rather stiff except for the case $T \ll 1$, we will need some special device to get such an enclosure.

Remark 5. In order to enclose the solution for (3.13), it seems that we have to estimate the value $F(t)$ pointwise in t and it is difficult in general. This problem is, for instance, resolved as follows. Equation (3.13) is translated to integral equation:

$$\boldsymbol{b} = \int_0^t P\boldsymbol{b}\, \mathrm{d}t + \int_0^t F(t)\, \mathrm{d}t.$$

Here, for example, if $F(t) = (\phi_j, (u_h + \alpha)^2)_\Omega$, then the term $(\phi_j, u_h^2)_\Omega$ can be estimated pointwise. The other term, e.g., $(\phi_j, \alpha^2)_\Omega$, is also enclosable by analogous techniques which will be described in the numerical example in Section 4. Furthermore, we do not need pointwise estimates for $F(t)$, if we use the standard L^2-Galerkin method on J. It is only required to estimate $F(t)$ in the $L^2(J)$ sense.

3.2. Fully discrete rounding

Now we present a verification technique using a simultaneous discretization of space and time.

In this subsection, we need an additional assumption for f in (2.1):

(A4) $f(\cdot, u)|_{t=0} \in H_0^1$ for any $u \in \tilde{H}$.

We introduce another finite-dimensional space $S^k \subset \tilde{H}^1 \equiv \{\phi \in H^1(J) \mid \phi(0) = 0\}$, $0 < k < 1$, on J with the following properties.

For any $v \in \tilde{H}^1 \cap H^2(J)$, there exist computable positive constants C_2 and C_3 such that

$$\inf_{\eta \in S^k} \| v - \eta \|_{L^2(J)} \leqslant \begin{cases} C_2 k^2 \| v'' \|_{L^2(J)}, \\ C_3 k \| v' \|_{L^2(J)}. \end{cases} \tag{3.14}$$

We adopt $S_{hk} = S_h \otimes S^k$ as the approximation space on $Q = \Omega \times J$, where S_h is the space defined in the previous subsection. For simplicity, suppose that $k = h$ and denote $S_x \equiv S_h$, $S_t \equiv S^h$ and $S_{xt} \equiv S_{hh}$ for fixed h.

Now, for $g \in L^2(J; L^2)$, the fully discrete rounding $u^h \equiv R(Ag) \in S_{xt}$ is

$$\int_0^T \left\{ (u_t^h, v)_\Omega + (\nabla u^h, \nabla v)_\Omega \right\} \mathrm{d}t = \int_0^T (g, v)_\Omega \, \mathrm{d}t, \quad v \in S_{xt}. \tag{3.15}$$

It is easily seen that there exists a unique u^h which satisfies (3.15).

Also we define projections $P_x : H_0^1(\Omega) \to S_x$ and $P_t : L^2(J) \to S_t$ by

$$(\nabla(\phi - P_x\phi), \nabla v)_\Omega = 0, \quad v \in S_x, \tag{3.16}$$

and

$$(\psi - P_t\psi, w)_J = 0, \quad w \in S_t, \tag{3.17}$$

respectively. Here, $(\cdot, \cdot)_J$ means the L^2 inner product on J. Then the proposition corresponding to Lemma 2 follows.

Lemma 6. *For $g \in H^1(J; L^2)$ with $g(x, 0) \in H_0^1$, let u and u^h be solutions of* (2.3) *and* (3.15), *respectively. Then, there exists a computable positive constant C such that*

$$\| \nabla(u - u^h) \| \leqslant Ch,$$

where $C = C(g, \| u_t^h \|)$.

Proof. First, setting $e = u - u^h$ by almost the same arguments as those of Lemma 2, we have, for each $v \in S_{xt}$,

$$\| \nabla e \|^2 \leqslant 2\big(\| e_t \| \, \| u - v \| + \tfrac{1}{2} \| \nabla (u - v) \|^2 \big). \tag{3.18}$$

Particularly, choosing $v = P_t P_x u$, from the definitions (3.1) and (3.14), it is implied that

$$\begin{aligned} \| u - P_t P_x u \| &\leqslant \| u - P_x u \| + \| u - P_t u \| + \| u - P_x u - P_t (u - P_x u) \| \\ &\leqslant 2 \| u - P_x u \| + \| u - P_t u \| \leqslant 2 C_1^2 h^2 \, | u |_{H^2(Q)} + C_2 h^2 \| u_{tt} \|. \end{aligned} \tag{3.19}$$

Similarly, we have

$$\begin{aligned} \| \nabla (u - P_t P_x u) \| &\leqslant 2 \| \nabla (u - P_x u) \| + \| \nabla u - P_t (\nabla u) \| \\ &\leqslant 2 C_1 h \, | u |_{H^2(Q)} + C_3 h \| \nabla u_t \|. \end{aligned} \tag{3.20}$$

Now we estimate the right-hand sides of (3.19) and (3.20). First, differentiating (2.3) with respect to time we obtain

$$(u_{tt}, \phi)_\Omega + (\nabla u_t, \nabla \phi)_\Omega = (g_t, \phi)_\Omega, \quad \phi \in H_0^1. \tag{3.21}$$

Setting $\phi = u_{tt}$ and integrating (3.21) with respect to time leads to

$$\int_0^t \| u_{tt} \|_\Omega^2 \, \mathrm{d}t + \| \nabla u_t(t) \|_\Omega^2 - \| \nabla u_t(0) \|_\Omega^2 \leqslant \int_0^t \| g_t \|_\Omega^2 \, \mathrm{d}t. \tag{3.22}$$

The smoothness of g in t implies that, in L^2 norm sense, $\lim_{t \to 0} u_t = g(\cdot, 0) \in H_0^1$. Therefore, noting that u_t is a solution of the parabolic problem

$$\begin{cases} \phi_t - \Delta \phi = g_t, & (x, t) \in \Omega \times J, \\ \phi(x, 0) = g(\cdot, 0), & x \in \Omega, \\ \phi(x, t) = 0, & (x, t) \in \partial\Omega \times J, \end{cases}$$

we have

$$\| \nabla u_t(0) \|_\Omega \leqslant \| \nabla g(\cdot, 0) \|_\Omega.$$

Thus by (3.22),

$$\| u_{tt} \|^2 \leqslant \| \nabla g(\cdot, 0) \|_\Omega^2 + \| g_t \|^2. \tag{3.23}$$

Further, setting $\phi = u_t$ in (3.21), integrating with respect to t from 0 to T and taking into account $u_t(0) = g(\cdot, 0)$, we get

$$\| \nabla u_t \|^2 \leqslant \tfrac{1}{2}\big(\| g(\cdot, 0) \|_\Omega^2 + \| g \|^2 + \| g_t \|^2 \big). \tag{3.24}$$

Here, we have used the estimates (3.6). Combining (3.18)–(3.24) with the estimates (3.6), (3.8) and

$$\| e_t \| \leqslant \| u_t^h \| + \| u_t \|,$$

we obtain the desired estimates with the constant

$$C^2 = 2\Big\{ K_1\big(C_1^2 K_2 + C_2 \sqrt{K_3} \big) + \tfrac{1}{2}\big(C_1 K_2 + C_3 \sqrt{K_4} \big)^2 \Big\}, \tag{3.25}$$

where

$$K_1 = \| g \| + \| u_t^h \|, \qquad K_2 = 4 \| g \|,$$
$$K_3 = \| \nabla g(\cdot, 0) \|_\Omega^2 + \| g_t \|^2, \qquad K_4 = \tfrac{1}{2}\left(\| g(\cdot, 0) \|_\Omega^2 + \| g \|^2 + \| g_t \|^2 \right). \quad \square \tag{3.26}$$

Thus, by the use of Lemma 6, we define the fully discrete rounding error $\mathrm{RE}(Ag)$ as

$$\mathrm{RE}(Ag) = \left\{ \phi \in \tilde{H} \mid \| \phi \|_{H_0} \leqslant hC \right\},$$

where C is the same constant determined by (3.25) and (3.26). Moreover, the definitions of $R(AG)$ and $\mathrm{RE}(AG)$ for the set of functions $G \subset H^1(J; L^2)$ with $G(\cdot, 0) \subset H_0^1$, are similar to those in Subsection 3.1.

Now let $\{\phi_j\}_{j=1,\ldots,M}$ be a basis for S_{xt} and let $\mathfrak{S}_1$ denote the set of all linear combinations of $\{\phi_j\}$ with interval coefficients. Then corresponding to Theorem 3, we have the following verification condition.

Theorem 7. *Let $U_h \in \mathfrak{S}_1$ and $\alpha, \beta \in \mathbb{R}^+$ and set $G = f(\cdot, U)$ for $U = (U_h, \alpha, \beta)$. If*

$$R(AG) \subset U_h, \tag{3.27}$$
$$\| \mathrm{RE}(AG) \|_{H_0} \leqslant \alpha, \tag{3.28}$$
$$\| G \| \leqslant \beta, \tag{3.29}$$

then there exists a solution $u \in \overline{U}$ for (2.5). Here, the triple (U_h, α, β) is defined in the previous subsection.

4. Verification procedure and numerical examples

In the present section, we describe a concrete algorithm for generation of the set which satisfies the verification conditions (3.27)–(3.29), and also give some numerical examples of verification.

We use an iterative procedure, which is similar to that in [13,15] etc., as follows. First, $u_0^h \in S_{xt}$, and $\alpha_0 \in \mathbb{R}^+$ are appropriately taken, normally, u_0^h will be chosen as a finite-element solution of (2.1) in S_{xt} and $\alpha_0 = 0$. Also we set $\beta_0 = \| (u_0^h)_t \|$ and $U_0 = (\{u_0^h\}, \alpha_0, \beta_0)$. Next, let $\epsilon_k > 0$, $1 \leqslant k \leqslant 3$, be small numbers. When $i \geqslant 1$, for

$$u_{i-1}^h = \sum_{j=1}^{M} \left[\underline{A}_j^{(i-1)}, \overline{A}_j^{(i-1)} \right] \phi_j,$$

set

$$\tilde{u}_{i-1}^h \equiv \sum_{j=1}^{M} \left[\underline{A}_j^{(i-1)} - \epsilon_1, \overline{A}_j^{(i-1)} + \epsilon_1 \right] \phi_j,$$

$$\tilde{\alpha}_{i-1} \equiv \alpha_{i-1} + \epsilon_2, \qquad \tilde{\beta}_{i-1} \equiv \beta_{i-1} + \epsilon_3,$$

and $\tilde{U}_{i-1} \equiv (\tilde{u}_{i-1}^h, \tilde{\alpha}_{i-1}, \tilde{\beta}_{i-1})$, which is the so-called ϵ-inflation. Then, we choose $u_i^h \in \mathfrak{S}_1$ and α_i, $\beta_i \in \mathbb{R}^+$ as

$$\left((u_i^h)_t, \phi_j\right) + \left(\nabla u_i^h, \nabla \phi_j\right) \supset \left(f(\tilde{U}_{i-1}), \phi_j\right), \quad 1 \leqslant j \leqslant M, \tag{4.1}$$

$$\alpha_i = hC\left(f(\tilde{U}_{i-1}), \| (u_i^h)_t \|\right), \tag{4.2}$$

and

$$\beta_i = \| f(\tilde{U}_{i-1}) \|, \tag{4.3}$$

respectively. Here, $f(\tilde{U}_{i-1}) \equiv f(\cdot, \tilde{U}_{i-1})$ and C is the constant of Lemma 6. Also (4.1) means that u_i^h is determined by an interval vector solution for the system of linear equations with interval right-hand side.

Then, by Theorem 7, the verification condition in computer can be written as in the next theorem.

Theorem 8. *Suppose that for some N*

$$u_N^h \subset \tilde{u}_{N-1}^h, \quad \alpha_N \leqslant \tilde{\alpha}_{N-1} \quad \text{and} \quad \beta_N \leqslant \tilde{\beta}_{N-1}. \tag{4.4}$$

Then there exixts a solution $u \in \overline{U}_N$ *for* (2.5), *where* $U_N = (u_N^h, \alpha_N, \beta_N)$.

Here, $u_N^h \subset \tilde{u}_{N-1}^h$ implies that each coefficient interval in u_N^h is included in the corresponding interval in $\tilde{u}_{N-1}^h$.

Remark 9. In order to compute the constant $C(f(\tilde{U}_{i-1}), \|(u_i^h)_t\|)$ in (4.2), we need the estimates for $\|(f(\cdot, u))_t\|$ and for $\|(f(\cdot, u))_x|_{t=0}\|_{\Omega}$ with $u \in U_{i-1}$. The former can be done, e.g., for $f(\cdot, u) = u^n$, as follows. Observe that $(u^n)_t = nu^{n-1}u_t$ and, using the a priori estimates for (2.3) (e.g., [12]), when $\Omega = (a, b)$, $\forall x \in \Omega$ and $\forall t \in J$, we have

$$\begin{aligned} |u(x, t)| &= \left| \int_a^x u_x(\xi, t)\, \mathrm{d}\xi \right| \leqslant \sup_{t \in J} d \| u_x(\cdot, t) \|_{L^2(\Omega)} \\ &\leqslant d \| f(\cdot, \tilde{U}_{i-2}) \| = d\beta_{i-1}, \end{aligned}$$

where $d = \sqrt{b-a}$. Thus we need an additional parameter γ_i corresponding to $\|u\|_{L^\infty(Q)}$, for $u \in U_i$, which is defined by $\gamma_i = d\beta_i$. But this quantity does not affect the verification condition. Hence, we can estimate the value $\|u\|_{L^\infty(Q)}$ for any $u \in \tilde{U}_{i-1}$ as $\|u\|_{L^\infty(Q)} \leqslant d(\beta_{i-1} + \epsilon_3) = d\tilde{\beta}_{i-1}$.

Therefore, we can iteratively calculate the desired estimates. And the latter is easily bounded by the assumption on $\tilde{H}$, using the similar estimates.

Now we will provide some numerical examples which were actually verified in computer by the procedure described above. We consider the following equation with interval coefficients:

$$\begin{cases} u_t - u_{xx} = pu^2 + ([q_1, q_2], [q_1', q_2']), & (x, t) \in I \times J, \\ u(x, 0) = 0, & x \in I, \\ u(0, t) = u(1, t) = 0, & t \in J, \end{cases} \tag{4.5}$$

where $I = J = (0, 1)$ and p is a given L^∞ function on I. Also $\mathfrak{Q} \equiv ([q_1, q_2], [q_1', q_2'])$ denotes the following set of functions:

$$\mathfrak{Q} = \{ q \in H_0^1(I) \mid q(x) \in [q_1, q_2] \text{ and } q'(x) \in [q_1', q_2'],\ x \in I \}.$$

Now let $\delta_x : 0 = x_0 < x_1 < \cdots < x_L = 1$ be a uniform partition of the interval I and set $I_i = (x_{i-1}, x_i)$ and $h = 1/L$. For simplicity, define the partition of J as $\delta_t = \delta_x$. When we denote by $\mathscr{M}^1$ the set of continuous piecewise linear functions on I or J, let $S_x = \{ v \in \mathscr{M}^1 \mid v(0) = v(1) = 0 \}$ and $S_t = \{ w \in \mathscr{M}^1 \mid w(0) = 0 \}$. Then, we adopt $S_{xt} = S_x \otimes S_t$ as the approximation space.

We have $\dim S_{xt} = L(L-1)$ and, by well-known results (e.g., [20]), the values of constants C_1–C_3 in the previous section can be taken as

$$C_1 = \frac{1}{\pi}, \quad C_2 = \frac{1}{\pi^2} \quad \text{and} \quad C_3 = \frac{1}{\pi}.$$

We also adopted the usual hat functions ϕ_j, as the basis of $\mathscr{M}^1$, which take the value 1 at x_j and 0 at x_k for $x_k \neq x_j$.

Further, in order to estimate u_i^h, α_i and β_i in (4.1)–(4.3), we need the calculations of integrals including the nonlinear term in $[\tilde{\alpha}_{i-1}]$. They are estimated with special attentions as follows (cf. [14]). For any $\alpha_{i-1} \in [\tilde{\alpha}_{i-1}]$, we have

$$\left(\alpha_{i-1}^2, \phi_j\right) \in [-1, 1] \| \alpha_{i-1} \|^2 \subset [-1, 1] \tilde{\alpha}_{i-1}^2.$$

The estimation of $\| \alpha_{i-1}^2 \|$ requires further consideration below. Let u and u_h be solutions for (2.3) and (3.15), respectively and set $e \equiv u - u_h$. Then observe that

$$\begin{aligned} \| e^2 \| &= \int_J \int_I e^4 \, dx \, dt \leqslant \int_J \left(\| e_x \|_{L^2(I)}^4 \int_I x^2 \, dx \right) dt \\ &\leqslant \tfrac{1}{3} \left(\int_J \| e_x \|^2 \, dt \right) \| e_x \|_{L^\infty(J;L^2)}^2 \\ &\leqslant \tfrac{1}{3} \| e_x \|^2 \left(\left\| (u^h)_x \right\|_{L^\infty(J;L^2)} + \| u_x \|_{L^\infty(J;L^2)} \right)^2. \end{aligned}$$

Furthermore, using the a priori estimates for the solution of (2.2) [12], we have

$$\| e^2 \|^2 \leqslant \tfrac{1}{3} \| e_x \|^2 \left(\left\| (u^h)_x \right\|_{L^\infty(J;L^2)} + \| g \| \right)^2. \tag{4.6}$$

Hence, we need an auxiliary parameter κ_i, corresponding to the quantity $\| \alpha_i^2 \|$, which is defined by

$$\kappa_i = \frac{1}{\sqrt{3}} \alpha_i \left(\left\| \left(u_i^h \right)_x \right\|_{L^\infty(J;L^2)} + \beta_i \right). \tag{4.7}$$

But from the right-hand side of (4.7), it is seen that this additional parameter has no effect on the verification condition, that is, if the conditions in Theorem 8 are satisfied for u_N^h, α_N and β_N, then so automatically for κ_N. Therefore, we may estimate $\| \alpha_{i-1}^2 \|$ as the right-hand side of (4.7) with α_i, u_i^h and β_i replaced by $\tilde{\alpha}_{i-1}$, $\tilde{u}_{i-1}^h$ and $\tilde{\beta}_{i-1}$, respectively.

We now illustrate numerical results of verified examples.

Case 1.
Equation: $u_t - u_{xx} = 0.1\, u^2 + ([0, 3], [0, 5])$, $(x, t) \in I \times J$.
Execution conditions: $L = 16$ ($\dim S_{xt} = 240$), $u_0^h = \alpha_0 = \beta_0 = 0$, $\epsilon_k = 10^{-2}$ $(k = 1, 2, 3)$.

Results: Iteration numbers: $N = 29$,
H^1-error bound: $\alpha = 0.597$,
L^2-bound of u_t: $\beta = 4.863$,
Coefficient intervals: $\min_{1 \leq j \leq M} \underline{A}_j = -1.028$, $\max_{1 \leq j \leq M} \overline{A}_j = 6.135$.

Case 2.
Equation: $u_t - u_{xx} = u^2 + ([0, 0.3184], [0, 1])$, $(x, t) \in I \times J$.
Execution conditions: $L = 22$ (dim $S_{xt} = 462$), $u_0^h = \alpha_0 = \beta_0 = 0$, $\epsilon_k = 10^{-2}$ $(k = 1, 2, 3)$.
Results: Iteration numbers: $N = 15$,
H^1-error bound: $\alpha = 0.0618$,
L^2-bound of u_t: $\beta = 0.580$,
Coefficient intervals: $\min_{1 \leq j \leq M} \underline{A}_j = -0.132$, $\max_{1 \leq j \leq M} \overline{A}_j = 0.813$.

Remark 10. In these numerical experiments, we used the usual floating-point number system with double precision. Therefore, the above results may include some unknown rounding errors in each step of the calculations. We need to use some arithmetic systems with guaranteed accuracy (e.g., [23]) for more rigorous verification, though the rounding errors caused by floating-point arithmetic are, in general, several orders of magnitude smaller than the truncation errors.

Remark 11. In the present situation, we implicitly have assumed that the map Af in (2.5) is retractive in a neighborhood of the solution which we seek. When this condition is not satisfied, we will have to devise another formulation, for example, based upon some Newton-like methods (e.g., [16]).

References

[1] O. Axelsson and V.A. Barker, *Finite Element Solution of Boundary Value Problems* (Academic Press, Orlando, 1984).
[2] P. Grisvard, *Elliptic Problems in Nonsmooth Domains* (Pitman, Berlin, 1983).
[3] E.W. Kaucher, Validating computation in a function space, in: R.E. Moore, Ed., *Reliability in Computing* (Academic Press, San Diego, 1988) 403–425.
[4] E.W. Kaucher and W.L. Miranker, *Self-validating Numerics for Function Space Problems* (Academic Press, New York, 1984).
[5] G. Kedem, A posteriori error bounds for two-point boundary value problems, *SIAM J. Numer. Anal.* **18** (1982) 431–448.
[6] U. Kulisch, The arithmetic of the digital computer: A new approach, *SIAM Rev.* **28** (1986) 1–40.
[7] U. Kulisch and W.L. Miranker, Eds., *A New Approach to Scientific Computation* (Academic Press, New York, 1983).
[8] U. Kulisch and H.J. Stetter, Eds., *Scientific Computation with Automatic Result Verification* (Springer, Wien, 1988).
[9] O.E. Lanford III, Computer assisted proofs in analysis, in: *Proc. Internat. Congress of Mathematicians*, Berkeley, CA, 1986 (Amer. Mathematical Soc., Providence, RI, 1987) 1385–1394.
[10] J.L. Lions and E. Magenes, *Non-homogeneous Boundary Value Problems and Applications* (Springer, Berlin, 1972).
[11] R.J. Lohner, Enclosing the solutions of ordinary initial and boundary value problems, in: E.W. Kaucher et al., Eds., *Computerarithmetic* (Teubner, Stuttgart, 1987).
[12] M. Luskin and R. Rannacher, On the smoothing property of the Galerkin method for parabolic equations, *SIAM J. Numer. Anal.* **19** (1981) 93–113.

[13] M.T. Nakao, A numerical approach to the proof of existence of solutions for elliptic problems, *Japan J. Appl. Math.* **5** (1988) 313–332.
[14] M.T. Nakao, A computational verification method of existence of solutions for nonlinear elliptic equations, in: M. Mimura and T. Nishida, Eds., *Recent Topics in Nonlinear PDE IV*, Lecture Notes Numer. Appl. Anal. **10**, North-Holland Math. Stud. **160** (Kinokuniya/North-Holland, Tokyo/Amsterdam, 1989) 101–120.
[15] M.T. Nakao, A numerical approach to the proof of existence of solutions for elliptic problems II, *Japan J. Appl. Math.* **5** (1990) 477–488.
[16] M.T. Nakao, A numerical verification method for the existence of weak solutions for nonlinear BVP, *J. Math. Anal. Appl.*, to appear.
[17] K. Nickel, Using interval methods for the numerical solution of ODE's, *Z. Angew. Math. Mech.* **66** (1986) 513–523.
[18] S.M. Rump, Solving nonlinear systems with least significant bit accuracy, *Computing* **29** (1982) 183–200.
[19] J. Schröder, A method for producing verified results for two-point boundary value problems, *Computing Suppl.* **6** (1988) 9–22.
[20] M.H. Schultz, *Spline Analysis* (Prentice-Hall, Englewood Cliffs, NJ, 1973).
[21] R. Temam, *Navier–Stokes Equations* (North-Holland, Amsterdam, 1977).
[22] E. Zeidler, *Nonlinear Functional Analysis and its Applications I* (Springer, New York, 1986).
[23] IBM High-accuracy arithmetic subroutine library (ACRITH), Program description and user's guide, SC 33-6164-02, 3rd edition, April 1986.